0590

REV. 9/00
For Other Than A Small Entity

MDhMORF-1 (as filed)
AEOMICA-1 (as amended herein)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	PENN *et al.*
Serial No.	:	09/774,203
Filed	:	January 29, 2001
For:	:	METHODS AND APPARATUS FOR PREDICTING, CONFIRMING, AND DISPLAYING FUNCTIONAL INFORMATION DERIVED FROM GENOMIC SEQUENCE
Group Art Unit	:	To Be Assigned

Hon. Commissioner
for Patents
Washington, D.C. 20231

TRANSMITTAL LETTER

Sir:

Transmitted herewith: [X] a Preliminary Amendment; [ ] a Reply to Office Action; [ ] a Supplemental Amendment; [ ] a substitute Specification; [ ] a Declaration; [ ] a Supplemental Declaration; [ ] a Power of Attorney; [ ] an Associate Power of Attorney; [ ] formal drawings; to be filed in the above-identified patent application.

FEE FOR ADDITIONAL CLAIMS

[ ] A fee for additional claims is not required.

[X] A fee for additional claims is required.

The additional fee has been calculated as shown below:

	CLAIMS REMAINING AFTER AMENDMENT		HIGHEST NUMBER PREVIOUSLY PAID FOR		PRESENT EXTRA		RATE	ADDITIONAL FEES
TOTAL CLAIMS	144	-	59	* =	85	X	$ 18 =	$1530.00
INDEPENDENT CLAIMS	11	-	3	** =	8	X	$ 80 =	$ 640.00
FIRST PRESENTATION OF A MULTIPLE DEPENDENT CLAIM							+ $270 =	$
							TOTAL	$2170

* If less than 20, insert 20.
** If less than 3, insert 3.

[X] A check in the amount of $$2170.00 in payment of the filing fee is transmitted herewith.

[X] The Director is hereby authorized to charge payment of any additional filing fees required under 37 C.F.R. § 1.16, in connection with the paper(s) transmitted herewith, or credit any overpayment of same, to deposit Account No. 06-1075. A duplicate copy of this transmittal letter is transmitted herewith.

[ ] Please charge $_____ to Deposit Account No. 06-1075 in payment of the filing fee. A duplicate copy of this transmittal letter is transmitted herewith.

EXTENSION FEE

[ ] The following extension is applicable to the Response filed herewith; [ ] $110.00 extension fee for response within first month pursuant to 37 C.F.R. § 1.136(a); [ ] $390.00 extension fee for response within second month pursuant to 37 C.F.R. § 1.136(a); [ ] $890.00 extension fee for response within third month pursuant to 37 C.F.R. § 1.136(a); [ ] $1,390.00 extension fee for response within fourth month pursuant to 37 C.F.R. § 1.136(a); $1,890.00 within fifth month pursuant to 37 C.F.R. § 1.136(a).

[ ] A check in the amount of [ ] $110.00; [ ] $390.00; [ ] $890.00; [ ] $1,390.00; [ ] $1,890.00 in payment of the extension fee is transmitted herewith.

[X] The Director is hereby authorized to charge payment of any additional fees required under 37 C.F.R. § 1.17 in connection with the paper(s) transmitted herewith, or to credit any overpayment of same, to Deposit Account No. 06-1075. A duplicate copy of this transmittal letter is transmitted herewith.

[ ] Please charge the [ ] $110.00; [ ] $390.00; [ ] $890.00; [ ] $1,390.00; [ ] $1,890.00; extension fee to Deposit Account No. 06-1075. A duplicate copy of this transmittal letter is transmitted herewith.

6 JULY 2001
Date

Daniel M. Becker
Registration No. 38,376
Attorney for Applicants

c/o FISH & NEAVE
Customer No. 1473
1251 Avenue of the Americas
New York, New York 10020-1104
Tel.: (650) 617-4000 (CA)

I hereby Certify that this Correspondence is being Deposited with the U.S. Postal Service as First Class Mail in an Envelope Addressed to: ASSISTANT COMMISSIONER FOR PATENTS WASHINGTON, D.C. 20231, ON
07-06-2001
LILY X. JIANG
Signature of Person Signing

Attorney Docket:
MDhMORF-1 (as filed)
AEOMICA-1 (as amended herein)

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants	:	Sharron G. Penn *et al.*
Serial No.	:	09/774,203
Confirm. No.	:	7320
Filed	:	January 29, 2001
For	:	METHODS AND APPARATUS FOR PREDICTING, CONFIRMING, AND DISPLAYING FUNCTIONAL INFORMATION DERIVED FROM GENOMIC SEQUENCE
Group Art Unit	:	To be assigned
Examiner	:	To be assigned

Palo Alto, CA
July 5, 2001

Hon. Commissioner for Patents
Washington, D.C. 20231

Second Preliminary Amendment Under 37 C.F.R. § 1.115

Sir:

Prior to examination of the above-identified application, please enter the following amendments and consider the following remarks.

IN PALM:

Please change the attorney docket number from "MDhMORF-1" to --AEOMICA-1-- .

IN THE SPECIFICATION:

(1) Please replace the paragraph that begins at line 19 of page 5 and ends at line 5 of page 6 with the following replacement paragraph:

Much of the coding sequence of the human genome is not homologous to known genes, making detection of open reading frames ("ORFs") and predictions of gene function difficult. Computational methods exist for predicting coding regions in eukaryotic genomes. Gene prediction programs such as GRAIL and GRAIL II, Uberbacher *et al.*, *Proc. Natl. Acad. Sci. USA* 88(24):11261-5 (1991); Xu *et al.*, *Genet. Eng.* 16:241-53 (1994); Uberbacher *et al.*, *Methods Enzymol.* 266:259-81 (1996); GENEFINDER, Solovyev *et al.*, *Nucl. Acids. Res.* 22:5156-63 (1994); Solovyev *et al.*, *Ismb* 5:294-302 (1997); and GENSCAN, Burge *et al.*, *J. Mol. Biol.* 268:78-94 (1997), predict many putative genes without known homology or function. Such programs are known, however, to give high false positive rates. Burset *et al.*, *Genomics* 34:353-367 (1996). Using a consensus obtained by a plurality of such programs is known to increase the reliability of calling exons from genomic sequence. Ansari-Lari *et al.*, Genome Res. 8(1):29-40 (1998).

(2) Please replace the paragraph that begins at line 30 of page 6 and ends at line 9 of page 7 with the following replacement paragraph:

It is common for microarrays to be derived from cDNA/EST libraries, either from those

previously described in the literature, such as those from the I.M.A.G.E. consortium, Lennon *et al.*, "The I.M.A.G.E. Consortium: an Integrated Molecular Analysis of Genomes and Their Expression, *Genomics* 33(1):151-2 (1996), or from the construction of "problem specific" libraries targeted at a particular biological question, R.S. Thomas *et al.*, *Toxicologist* 54:68-69 (2000). Such microarrays by definition can measure expression only of those genes found in EST libraries, and thus have not been useful as probes for genes discovered solely by genomic sequencing.

(3) Please replace the paragraph that begins at line 33 of page 23 and ends at line 8 of page 25 with the following replacement paragraph:

As discussed below, and further described in detail in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668; PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entirety, we have used the methods and apparatus of the present invention to identify more than 15,000 exons in human genomic sequence whose expression we have confirmed in at least one human tissue or cell

type. Fully two-thirds of the exons belong to genes that were not at the time of our discovery represented in existing public expression (EST, cDNA) databases, making the methods and apparatus of the present invention extremely powerful tools for novel gene discovery.

(4) Please replace the paragraph that begins at line 28 of page 76 and ends at line 3 of page 77 with the following replacement paragraph:

One third of the probe sequences (as amplified) produced an exact match (BLAST Expect ("*E*") values less than 1e-100 (1×10^{-100})) to either an EST (20% of sequences) or a known mRNA (13% of sequences). A further 22% of the probe sequences showed some homology to a known EST or mRNA (BLAST *E* values from 1e-5 (1×10^{-5}) to 1e-99 (1×10^{-99})). The remaining 45% of the probe sequences showed no significant sequence homology to any expressed, or potentially expressed, sequences present in public databases.

(5) Please replace the paragraph that begins at line 26 of page 80 and ends at line 10 of page 81 with the following replacement paragraph:

FIG. 7A is a matrix presenting the expression of all verified sequences that showed signal intensity greater than 3 in at least one tissue. Each clone is represented by a column in the matrix. Each of the 10 tissues assayed is represented by a separate row in the matrix, and relative expression (expression ratio) of a clone

in that tissue is indicated at the respective node by intensity of green shading, with the intensity legend shown in panel B. The top row of the matrix ("EST Hit") contains "bioinformatic" rather than "physical" expression data – that is, presents the results returned by query of EST, NR and SwissProt databases using the probe sequence. The legend for "bioinformatic expression" (*i.e.*, degree of homology returned) is presented in panel C. Briefly, white is known, black is novel, with gray depicting nonidentical with significant homology (white: *E* values < 1e-100 (1×10^{-100}); gray: *E* values from 1e-5 (1×10^{-5}) to 1e-99 (1×10^{-99}); black: *E* values > 1e-5 (1×10^{-5})).

(6) Please replace the paragraph at lines 7 - 16 of page 82 with the following replacement paragraph:

FIG. 8 shows in dashed line the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect ("*E*") value of greater than 1e-30 (1×10^{-30}) (designated "unknown") upon query of existing EST, NR and SwissProt databases, and shows in solid line the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect value of less than 1e-30 (1×10^{-30}) ("known"). Note that biological background noise has an averaged normalized Cy3 signal intensity of 0.2.

(7) Please replace the paragraph that begins at line 18 of page 96 and ends at line 22 of page 97 with the following replacement paragraph:

Using this threshold, we identified over 15,000 single exon probes that produce significant signal in one or more of ten tested tissues/cell types. The exact structures of these single exon probes are clearly presented in the SEQUENCE LISTINGs included in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668; PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entireties.

(8) Please replace the paragraph that begins at line 12 of page 98 and ends at line 13 of page 99 with the following replacement paragraph:

The exact structures of these single exon probes are clearly presented in the SEQUENCE LISTINGs included in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668;

PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entireties.

IN THE CLAIMS:

Please add the following new claims 21 - 92:

21 (new). A high throughput, microarray-based method to confirm predicted exons, comprising:

detecting hybridization by transcript-derived nucleic acids to microarray probes that include genomic sequence predicted to contribute to no more than one exon,

detectable hybridization confirming the prediction of the exon included in each of said detectably hybridized probes.

22 (new). The method of claim 21, wherein at least 75% of the probes of said microarray include genomic sequence predicted to contribute to no more than one exon.

23 (new). The method of claim 21, wherein at least 90% of the probes of said microarray include genomic sequence predicted to contribute to no more than one exon

24 (new). The method of claim 21, wherein at least 95% of the probes of said microarray include genomic sequence predicted to contribute to no more than one exon.

25 (new). The method of claim 21, wherein said genomic sequence is human genomic sequence.

26 (new). The method of claim 21, wherein said prediction is output from a computer program selected from the group consisting of GenScan, Diction, Genefinder, and Grail.

27 (new). The method of claim 26, wherein said prediction is output from GenScan.

28 (new). The method of claim 21, wherein said microarray has probes that collectively include exons predicted from all chromosomes of a eukaryotic organism.

29 (new). The method of claim 28, wherein said eukaryotic organism is a human being.

30 (new). The method of claim 21, wherein said microarray has probes that include exons predicted from human chromosome 22.

31 (new). The method of claim 21, wherein each of said predicted exons is represented by a plurality of probes on said array.

32 (new). The method of claim 21, wherein said microarray includes between 5,000 and 19,000 probes.

33 (new). The method of claim 21, wherein the genomic sequence included within said probes is selected at least in part based upon considerations of base composition and/or hybridization binding stringency.

34 (new). The method of claim 21, wherein said probes include at least 50 nt of predicted exon.

35 (new). The method of claim 21, wherein said probes include at least 75 nt of predicted exon.

36 (new). The method of claim 21, wherein said probes are amplified from genomic DNA.

37 (new). The method of claim 21, wherein said probes are chemically synthesized.

38 (new). The method of claim 21, wherein said probes are noncovalently attached to the substrate of said microarray.

39 (new). The method of claim 21, wherein said probes are covalently attached to the substrate of said microarray.

40 (new). The method of claim 21, wherein said probes are disposed on said microarray substrate by ink jet.

41 (new). The method of claim 21, wherein the substrate of said microarray is a glass slide.

42 (new). The method of claim 21, further comprising the antecedent step of:

contacting said microarray with at least a first sample of transcript-derived nucleic acids, said nucleic acids being detectably labeled.

43 (new). The method of claim 42, wherein said transcript-derived nucleic acids are first strand cDNA.

44 (new). The method of claim 43, wherein said cDNAs are fluorescently labeled.

45 (new). The method of claim 44, wherein said fluorescent label is selected from the group consisting of Cy3 and Cy5.

46 (new). The method of claim 42, wherein said contacting step comprises contacting said microarray concurrently with a first sample of transcript-derived nucleic acids and with a second sample of transcript-derived nucleic acids, wherein said first and second samples are labeled respectively with a first and a second label, said first and second labels being separately detectable.

47 (new). The method of claim 46, wherein said detecting includes normalizing and background correcting signals from each of said labels.

48 (new). The method of claim 46, wherein said labels are Cy3 and Cy5.

49 (new). The method of claim 46, wherein said first sample includes transcript-derived nucleic acids pooled from a plurality of tissues and/or cell types.

50 (new). The method of claim 49, wherein said pool includes transcript-derived nucleic acids from a plurality of human cell lines.

51 (new). The method of claim 49, wherein the transcript-derived nucleic acids of said second sample are derived from a cell line or normal tissue.

52 (new). The method of claim 51, wherein the transcript-derived nucleic acids of said second sample are derived from a source within the group of human tissues and

cell lines consisting of: brain, heart, liver, fetal liver, placenta, lung, bone marrow, HeLa cells, BT474 cells and HBL 100 cells.

53 (new). A method of identifying potential false positive exon predictions, comprising:

detecting hybridization by transcript-derived nucleic acids to a microarray that has probes that include genomic sequence predicted to contribute to no more than one exon,

absence of detectable hybridization identifying as a potential false positive the exon predicted in each undetectably hybridized probe.

54 (new). A method of identifying one or more genes expressed by one or more eukaryotic cells having a genome that averages at least one intron per gene, comprising:

(a) contacting a cDNA sample prepared by enzymatically copying messenger RNA obtained from said eukaryotic cell(s) into cDNA, wherein said cDNA comprises a detectable label, with a plurality of single exon probes, each said single exon probe comprising a discrete nucleic acid sequence encoding all or a portion of a single exon of said eukaryotic genome that specifically hybridizes at high stringency to a target nucleic acid when said target nucleic acid is present in said cDNA sample;

(b) detecting a signal from each said single exon probe that is specifically hybridized to said target nucleic acid, wherein the presence of said signal indicates the expression of a gene comprising said single exon by said eukaryotic cell(s).

55 (new). A method of identifying one or more genes expressed by one or more human cells, comprising:

(a) contacting a cDNA sample prepared by copying messenger RNA obtained from said human cell(s) into cDNA using reverse transcriptase, wherein said cDNA comprises a detectable label, with a nucleic acid microarray, said microarray comprising a substantially planar glass substrate comprising (i) at least 5000 addressable locations to which single exon probes are bound, each said single exon probe comprising a discrete nucleic acid sequence encoding all or a portion of a single exon of a human genome that is specifically hybridizable at high stringency to a target nucleic acid, wherein said target nucleic acid is a sequence encoding all or a portion of an expressed gene, or a complementary sequence thereof, and (ii) one or more additional locations to which control nucleic acid sequences are bound; and

(b) generating a signal from each said addressable location, wherein the presence of a signal at a specific addressable location indicates the expression by said human cell(s) of a gene comprising the single exon probe bound to that addressable location.

56 (new). A high throughput, microarray-based method of grouping exons into a common gene, comprising:

comparing the patterns of tissue and/or cell-type expression of exons predicted from a contiguous region of genomic DNA,

wherein said patterns of expression have been determined by detecting hybridization of transcript-derived nucleic acids from a plurality of tissues and/or cell types to microarray probes, each of said probes including genomic sequence predicted to contribute to no more than one of said exons, said microarray including probes that collectively comprise all of said exons,

consensus in said expression patterns identifying exons that are groupable into a common gene.

57 (new). The method of claim 56, wherein said gene is a human gene.

58 (new). The method of claim 56, wherein said patterns are detected by detecting (i) fluorescence intensity, (ii) the ratio of intensity as between concurrently hybridized first and second samples, or (iii) a combination of (i) and (ii).

59 (new). A nucleic acid microarray comprising:

a substrate comprising a plurality of addressable locations to which nucleic acid sequences are bound; and

a plurality of single exon probes bound at said addressable locations, each said single exon probe comprising a discrete nucleic acid sequence encoding all or a portion of a single exon of a eukaryotic genome averaging at least one intron per gene that is specifically hybridizable at high stringency to a target nucleic acid, wherein said target nucleic acid is a sequence encoding all or a portion of an expressed gene, or a complementary sequence thereof.

60 (new). A nucleic acid microarray comprising:

a substantially planar glass substrate comprising (i) at least 5000 addressable locations to which single exon probes are bound, each said single exon probe comprising a discrete nucleic acid sequence encoding all or a portion of a single exon of a human genome that is specifically hybridizable at high stringency to a target nucleic acid, wherein said target nucleic acid is a sequence encoding all or a portion of an expressed gene, or a complementary

sequence thereof, and (ii) one or more additional locations to which control nucleic acid sequences are bound.

61 (new). A single exon nucleic acid microarray, comprising:

a plurality of nucleic acid probes addressably disposed upon a substrate,

wherein at least 50% of said probes include genomic sequence predicted to contribute to no more than one exon of a eukaryotic genome, said eukaryotic genome averaging at least one intron per gene, and wherein said plurality of nucleic acid probes averages at least 50 nt in length.

62 (new). The microarray of claim 61, wherein at least 75% of said nucleic acid probes include genomic sequence predicted to contribute to no more than one exon of a eukaryotic genome.

63 (new). The microarray of claim 61, wherein at least 90% of the probes of said microarray include genomic sequence predicted to contribute to no more than one exon of a eukaryotic genome.

64 (new). The microarray of claim 61, wherein at least 95% of the probes of said microarray include genomic sequence predicted to contribute to no more than one exon of a eukaryotic genome.

65 (new). The microarray of claim 61, wherein said microarray has probes that collectively include exons predicted from all chromosomes of a eukaryotic genome.

66 (new). The microarray of claim 61, wherein said eukaryotic genome is a human genome.

67 (new). The microarray of claim 65, wherein said eukaryotic genome is a human genome.

68 (new). The microarray of claim 61, wherein said prediction is output from a computer program selected from the group consisting of GenScan, Diction, Genefinder, and Grail.

69 (new). The microarray of claim 68, wherein said prediction is output from GenScan.

70 (new). The microarray of claim 61, wherein each of said predicted exons is represented by a plurality of probes on said array.

71 (new). The microarray of claim 61, wherein said microarray includes between 5,000 and 19,000 probes.

72 (new). The microarray of claim 61, wherein the genomic sequence included within said probes is selected at least in part based upon considerations of base composition and/or hybridization binding stringency.

73 (new). The microarray of claim 61, wherein said probes have been amplified from genomic DNA.

74 (new). The microarray of claim 61, wherein said probes have been chemically synthesized.

75 (new). The microarray of claim 61, wherein said probes are noncovalently attached to the substrate of said microarray.

76 (new). The microarray of claim 61, wherein said probes are covalently attached to the substrate of said microarray.

77 (new). The microarray of claim 61, wherein said probes are disposed on said microarray substrate by ink jet.

78 (new). The microarray of claim 61, wherein said substrate is a glass slide.

79 (new). The microarray of claim 61, wherein each of said probes is disposed on said array with its reverse complement.

80 (new). The microarray of claim 61, further comprising control probes.

81 (new). The microarray of claim 61, wherein at least 50% of said exon-including nucleic acid probes comprise, contiguous to a first end of said predicted exon, a first intronic and/or intergenic sequence that is identically contiguous to said exon in the human genome, and further comprise, contiguous to a second end of said predicted exon, a second intronic and/or intergenic sequence that is identically contiguous to said exon in the human genome.

82 (new). A software data structure for annotating nucleic acid sequence with confirmed bioinformatic predictions, the data structure stored in a machine readable medium and comprising:

a plurality of sequence entries, each sequence entry including (i) a sequence identifier and (ii) software means for relating said sequence identifier to data that

encode a confirmed prediction of a biological function of the nucleic acid sequence identified by said sequence identifier.

83 (new). The software data structure of claim 82, wherein said confirmed biological function is contribution to a mature mRNA transcript.

84 (new). The software data structure of claim 83, wherein said prediction is output from GenScan.

85 (new). The software data structure of claim 83, wherein said prediction has been confirmed by the method of claim 21.

86 (new). The software data structure of claim 82, wherein said software relating means is the common inclusion of said confirmed prediction data in a single record with said sequence identifier.

87 (new). The software data structure of claim 82, wherein said software relating means links said sequence identifier to confirmed prediction data present in a distinct record.

88 (new). The software data structure of claim 82, wherein said sequence entries further comprise:

software means for relating said sequence identifier to data that encode at least one nucleic acid sequence identified by said identifier.

89 (new). The software data structure of claim 88, wherein said sequence entries further comprise:

software means for relating said sequence identifier and/or said at least one nucleic acid sequence to

data that encode a measure of similarity of the at least one nucleic acid sequence to at least one nucleic acid sequence prior-accessioned into a database.

90 (new). The software data structure of claim 89, wherein said sequence entries further comprise:

software means for relating said sequence identifier and/or said at least one nucleic acid sequence to data that encode a textual description of said at least one similar prior-accessioned nucleic acid sequence.

91 (new). The software data structure of claim 82, wherein said sequence entries further comprise:

software means for relating said sequence identifier to data that encode a chromosomal map location of the sequence identified by said sequence identifier.

92 (new). An isolated nucleic acid having exons that have been commonly grouped by the method of claim 56.

REMARKS

Amendments to the Specification

(1) Applicants have amended the specification by replacing the paragraph that begins at line 19 on page 5 and ends at line 5 of page 6 in order to correct a clerical error in the spelling of the gene finding program GENSCAN (from "GENESCAN"). The correct spelling appears in the specification as filed at page 25, line 28; page 27, line 23; p. 65, lines 30 - 31; and p. 94, line 5; accordingly, no new matter has been added.

(2) Applicants have amended the specification by replacing the paragraph that begins at line 30 of page 6 and ends at line 9 of page 7 in order to provide the full citation for a reference originally disclosed as "in press". The full citation can be found in the specification as filed at p. 39, lines 4 - 5; accordingly, no new matter has been added.

(3) Applicants have amended the specification by replacing the paragraph that begins at line 33 of page 23 and ends at line 8 of page 25 (i) to bring the serial number of applicants' GB priority application into conformity with the practice of the respective foreign patent office, adding the required prefatory two digit year designator to the serial number as originally set forth in the specification of the instant application as filed, and (ii) to replace the docket number, filing date, and title, for each of applicants' ten international priority applications with their respective application serial numbers. No new matter has been added.

(4) Applicants have amended the specification by replacing the paragraph that begins at line 28 of page 76 and ends at line 3 of page 77 in order to adjust the font, thus clarifying that the lower case "e" in the BLAST expect ("*E*") values signals base 10 exponentiation. The intent of the original font would readily have been understood by the skilled artisan familiar with BLAST, and is in accord with the parenthetical equivalences set forth in the specification as filed at page 11, lines 4 - 7; accordingly, no new matter has been added.

(5) Applicants have amended the specification by replacing the paragraph that begins at line 26 of page 80 and ends at line 10 of page 81 to adjust the font, thus

clarifying that the lower case "e" in the BLAST expect ("*E*") values signals base 10 exponentiation. The intent of the original font would readily have been understood by the skilled artisan familiar with BLAST, and is in accord with the parenthetical equivalences set forth in the specification as filed at page 11, lines 4 - 7; accordingly, no new matter has been added.

(6) Applicants have amended the specification by replacing the paragraph at lines 7 - 16 of page 82 to adjust the font, thus clarifying that the lower case "e" in the BLAST expect ("*E*") values signals base 10 exponentiation. The intent of the original font would readily have been understood by the skilled artisan familiar with BLAST, and is in accord with the parenthetical equivalences set forth in the specification as filed at page 11, lines 4 - 7; accordingly, no new matter has been added. The applicants have further amended the paragraph to replace reference to a blue line with reference to a solid line, a change occasioned by rendering of the color figure in black and white. The reference is unambiguous, inasmuch as (i) there are only two lines, the other of which is correctly identified, and (ii) the line is identified as solid in the specification as filed at page 11, line 5.

(7) Applicants have amended the specification by replacing the paragraph that begins at line 18 of page 96 and ends at line 22 of page 97 (i) to bring the serial number of applicants' GB priority application into conformity with the practice of the respective foreign patent office, adding the required prefatory two digit year designator to the serial number as originally set forth in the specification of the instant application as filed, and (ii) to replace the docket number, filing date, and title, for each of applicants' ten

international priority applications with their respective application serial numbers. No new matter has been added.

(8) Applicants have amended the specification by replacing the paragraph that begins at line 12 of page 98 and ends at line 13 of page 99 (i) to bring the serial number of applicants' GB priority application into conformity with the practice of the respective foreign patent office, adding the required prefatory two digit year designator to the serial number as originally set forth in the specification of the instant application as filed, and (ii) to replace the docket number, filing date, and title, for each of applicants' ten international priority applications with their respective application serial numbers. No new matter has been added.

Amendments to the Claims

Applicants herein add new claims 21 - 92 more particularly to point out and distinctly claim applicants' invention. Support for new claims 21 - 92 can be found throughout the specification, and particularly as set forth below. No new matter has been added.

Claims 21 - 53

Claims 21 - 52 are drawn to a high throughput microarray-based method to confirm predicted exons, comprising detecting hybridization by transcript-derived nucleic acids to microarray probes that include genomic sequence predicted to contribute to no more than one exon, whereby detectable hybridization confirms the prediction of the exon included in each of the detectably hybridized probes. **Claim 53** is drawn to the converse, a method of identifying potential false positive exon predictions,

whereby absence of detectable hybridization in the method of claim 21 identifies as a potential false positive the exon predicted in each undetectably hybridized probe. Support for the methods can be found throughout the specification and particularly between page 30, line 14 and page 60, line 18; in Examples 2 and 5; and in claims 15 - 18 as originally filed. Additional support for elements of dependent **claims 22 - 52** can be found throughput the specification and particularly as follows.

Support for the frequencies recited in **claims 22 - 24** can be found particularly at page 44, lines 9 - 20.

Support for application of the methods of the present invention to human genomic sequence, as recited in **claim 25,** is found throughout the specification and particularly at p. 61, lines 24 - 32; Examples 1, 2 and 5 with associated FIGS. 5 - 11; and at p. 18, lines 25 - 31; p. 20, lines 17 - 22; p. 24, line 33 - p. 25, line 8; p. 26, lines 14 - 22; p. 31, lines 6 - 16; p. 33, lines 7 - 10; p. 35, lines 6 - 9; p. 39, lines 11 - 19; and p. 55, lines 10 - 17. The aforedescribed locations are exemplary, not exhaustive.

Support for output of predictions from GenScan, Diction, Genefinder, and Grail, as recited in **claim 26 and claim 27** can be found particularly in Examples 1, 2, and 5; p. 25, lines 24 - 28; p. 27, lines 21 - 24; p. 65, lines 28 - 33; and p. 94, lines 4 - 6.

Collective inclusion of exons predicted from all chromosomes of a eukaryotic organism, particularly a human being, as recited in **claims 28 and 29,** is particularly exemplified in Examples 1, 2 and 5.

Inclusion of probes that include exons predicted from human chromosome 22, as recited in **claim 30,** can be

found in Table 2's inclusion of AC006548-9, AC006548 mapping to homo sapiens chromosome 22q11.

Support for representation of each predicted exon by a plurality of probes, as recited in **claim 31**, can be found particularly at p. 38, lines 12 - 16.

Support for including between 5,000 and 19,000 probes on the microarray, as recited in **claim 32**, can be found particularly at p. 38, lines 7 - 9.

Support for **claim 33** can be found particularly at p. 29, line 19 to p. 30, line 2.

Support for the probe sizes recited in **claims 34 - 35** can be found particularly at p. 54, lines 5 - 14; p. 31, lines 8 - 16.

Support for amplification of probes from genomic DNA, as set forth in **claim 36**, can be found throughout the specification and particularly at p. 30, line 27 - p. 35, line 25; and Examples 1, 2, 4 and 5.

Support for chemical synthesis of probes, as recited in **claim 37**, can be found particularly at p. 32, lines 9 - 12; and p. 40, lines 15 - 18;

Noncovalent attachment of probes to the microarray substrate, as recited in **claim 38**, is exemplified in Examples 2 and 5, and is further particularly described at p. 36, lines 11 - 15; p. 46, lines 22 - 27. Covalent attachment of probes to the microarray substrate, as recited in **claim 39**, is particularly recited at p. 36, lines 11 - 12.

Use of ink jet technology for creation of the microarray, as recited in **claim 40**, is described particularly at p. 36, lines 24 - 26.

Use of glass slides as the microarray substrate, as set forth in **claim 41**, is exemplified in Examples 2 and 5, and also particularly described at p. 31, line 16; and p. 35, lines 26 - 27.

The additional, antecedent, method step recited in **claim 42** -- "contacting said microarray with at least a first sample of transcript-derived nucleic acids, said nucleic acids being detectably labeled" - is particularly described in Example 2; and p. 48, line 13 - p. 50, line 14. Use of labeled, particularly fluorescently labeled, first strand cDNA, as recited in **claims 43 and 44**, is particularly exemplified in Examples 2 and 5.

Particular support for labeling the transcript-derived nucleic acids, particularly first strand cDNAs, with Cy3 or Cy5, as recited in **claims 45 and 48**, can be found particularly in Examples 2 and 5, and additionally at p. 10, line 31; p. 50, lines 2 - 3; and p. 69, lines 18 - 20.

Concurrent two-color microarray hybridization, as recited in **claim 46**, is exemplified in Examples 2 and 5, and is further particularly described at p. 10, lines 7 - 12; and p. 69, lines 9 - 13.

Normalization of signals, as recited in **claim 47**, can be found particularly at p. 10, lines 31 - 33; p. 69, lines 28 - 31; Example 2 (particularly at p. 79, lines 28 - 32 and p. 82, lines 2 - 6); Example 5 (particularly at p. 96, lines 10 - 17). Background correction of signals, as further recited in **claim 47**, can be found particularly between p. 95, line 12, and page 96, line 9.

Use of a pooled nucleic acid control, particularly a pool that includes transcript-derived nucleic acids from a plurality of human cell lines, as recited in **claims 49 - 50**, can be found particularly at p. 49, lines 2 - 9; p. 78, lines 4 - 15.

Use of transcript-derived nucleic acids derived from cell lines and normal tissues, and particularly derived from a source within the group of human tissues and cell lines consisting of brain, heart, liver, fetal liver, placenta, lung, bone marrow, HeLa cells, BT474 cells and HBL

100 cells, as recited in **claims 51 and 52**, is described particularly in Examples 2 and 5.

Claims 54 - 55

Claims 54 and 55 are drawn to methods of identifying genes expressed by cells of a eukaryote averaging at least one intron per gene (**claim 54**), in particular human cells (**claim 55**), by hybridization of cDNA samples to single exon microarrays of the present invention, whereby the presence of a signal indicates the expression in the cell(s) from which the cDNA derives of a gene that comprises the exon of the signaling probe. Support can be found throughout the specification, particularly as set forth above for claims 21 - 53.

Claims 56 - 58

Claims 56 - 58 are drawn to a high throughput, microarray-based method of grouping exons into a common gene based upon coregulated expression of exons predicted from a contiguous region of genomic DNA. Support for **claim 56** can be found particularly at p. 70, line 32 - p. 71, line 7.

Claims 59 - 81

Claims 59 - 81, which correspond to original claims 1 - 14, are drawn to single exon microarrays. Support for such microarrays is found throughout the specification, particularly including Examples 2, 4, and 5.

Additional support for various elements of **claim 59** can be found as follows. Exemplary high stringency conditions for microarray hybridization are described, *inter alia*, at page 54, lines 25 - 31; and p. 79, lines 10 - 16.

Particular support for exon microarrays from eukaryotic genomes having at least one intron per gene can be found at p. 48, lines 3 - 9 and in claim 1 as filed.

Additional support for various elements of **claim 60** can be found as follows. Support for microarrays having at least 5000 addressable probes can be found particularly at p. 38, lines 1 - 5. Support for inclusion of control probes can be found, *inter alia*, at p. 36, line 27 - p. 37, line 2; p. 76, lines 21 - 24; p. 95, line 12 - p. 96, line 9.

Additional support for various elements of **claim 61** can be found as follows. Support for the frequency of probes that have genomic sequence predicted to contribute to no more than a single exon can be found, *inter alia*, particularly at page 44, lines 9 - 20 and claim 1 as filed. Support for the average number of introns per gene can be found particularly at p. 47, line 12 - p. 48, line 9. Support for a minimum probe length of 50 nt can be found particularly at p. 54, lines 5 - 14.

Additional support for the frequencies recited in **claims 62 - 64** can be found particularly at page 44, lines 9 - 20.

Support for collective inclusion of exons predicted from all chromosomes of a eukaryotic organism, particularly a human being, as recited in **claims 65 - 67**, is particularly exemplified in Examples 1, 2 and 5.

Support for output of predictions from GenScan, Diction, Genefinder, and Grail, as recited in **claims 68 - 69** can be found particularly in Examples 1, 2, and 5; p. 25, lines 24 - 28; p. 27, lines 21 - 24; p. 65, lines 28 - 33; and p. 94, lines 4 - 6.

Support for representation of each predicted exon by a plurality of probes, as recited in **claim 70**, can be found particularly at p. 38, lines 12 - 16.

Support for including between 5,000 and 19,000 probes on the microarray, as recited in **claim 71**, can be found particularly at p. 38, lines 7 - 9.

Support for **claim 72** can be found particularly at p. 29, line 19 to p. 30, line 2

Support for amplification of probes of the microarray from genomic DNA, recited in **claim 73**, can be particularly be found exemplified in Examples 2, 4, and 5; and at p. 30, line 27 - p. 35, line 25.

Chemical synthesis of the probes of the microarray, as recited in **claim 74**, is particularly described in the specification at p. 32, lines 9 - 12; and p. 40, lines 15 - 18.

Noncovalent attachment of probes to the microarray substrate, as set forth in **claim 75**, is exemplified in Examples 2 and 5, and is further particularly described at p. 36, lines 11 - 15; p. 46, lines 22 - 27. Covalent attachment of probes to the microarray substrate, as recited in **claim 76**, is particularly recited at p. 36, lines 11 - 12.

Use of ink jet technology for creation of the microarray, as recited in **claim 77**, is described particularly at p. 36, lines 24 - 26.

Use of glass slides as the microarray substrate, as recited in **claim 78**, is exemplified in Examples 2 and 5, and also particularly described at p. 31, line 16; and p. 35, lines 26 - 27.

Claim 79, which recites that each of the probes is disposed on the array with its reverse complement, is supported throughout the specification's description of use of double-stranded amplified nucleic acids as probes, which double-stranded nucleic acids necessarily include the reverse complement of each of the two strands of the duplex. Further support can be found at p. 55, line 27 to p. 56, line 3.

Use of control probes, as set forth in **claim 80**, is exemplified in Examples 2, 4, and 5, and is further described, *inter alia*, at p. 36, lines 27 - 32.

Claim 81 finds particular support in claim 5 as originally filed, and specification p. 32, lines 4 - 8; p. 33, lines 3 - 6; page 55, lines 10 - 17; and p. 56, lines 4 - 10.

Claims 82 - 91

Claims 82 - 91 are drawn to software data structures for annotating nucleic acid sequence with confirmed bioinformatic predictions. Support can be found throughout the specification, and particularly in FIG. 1; in the specification from page 16, line 33 to page 17, line 15; p. 60, line 19 - p. 61, line 23 (notably on p. 61, lines 9 - 23); from page 14, line 12 to p. 23, line 32 (particularly at p. 15, lines 1 - 15); in the data structure output embodiments as shown in Tables 1, 2, and 3; and in the data structure output embodiments (visual displays) described from line 1 of page 62 to line 17 of p. 72 and shown in FIGS. 3 - 4 and 9 - 10.

Claims 83 - 85 particularly recite that the confirmed biological function is contribution of the genomic sequence to a mature mRNA transcript (*i.e.*, expression). Support can be found throughout the specification and particularly between page 30, line 14 and page 60, line 18; in the data structure output embodiments shown in FIGS. 3 - 4 and 9 - 10; and in Examples 2 and 5.

Common inclusion of confirmed prediction data in a single record with the sequence identifier, as recited in **claim 86**, is particularly described at p. 61, lines 15 - 17. Relating confirmed prediction data to the sequence identifier

by linking to a distinct record, as recited in **claim 87**, is particularly described at p. 61, lines 18 - 20.

Inclusion of sequence data itself, as recited in **claim 88**, is particularly described at p. 61, lines 13 - 15.

Support for inclusion within the sequence entries of software means for relating the sequence identifier to a measure of similarity to prior-accessioned sequences, as recited in **claim 89**, can be found particularly at p. 68, lines 13 - 33; and in the output embodiments shown in FIGS. 4, 8 and 9.

Support for inclusion in the sequence entries of software means for relating the sequence identifier to data that encode a textual description of similar, prior-accessioned nucleic acid sequences, as recited in **claim 90**, is particularly shown in the output embodiments of Table 2.

Inclusion in the sequence entries of software means for relating sequence identifiers to chromosomal map locations, as recited in **claim 91**, is found particularly in the description of the visual display output embodiments, particularly at p. 63, lines 21 - 25.

Claim 92

Claim 92 is drawn to the product of the process of claim 56, an isolated nucleic acid having exons that have been commonly grouped by the referenced method. Support can be found as for claim 56.

Respectfully submitted,

Daniel M. Becker (Reg. No. 38,376)
Attorney for Applicants

FISH & NEAVE
Customer No. 1473
1251 Avenue of the Americas
New York, NY 10020
650.617.4000
212.596.9090 (FAX)

Attachment:

- Appendix Pursuant to 37 C.F.R. § 1.121(b)(1)(iii)

Appendix Pursuant to 37 C.F.R. § 1.121(b)(1)(iii)

(1) Herewith, a marked up copy of the paragraph replacing that which begins at line 19 on page 5 and ends at line 5 of page 6:

> Much of the coding sequence of the human genome is not homologous to known genes, making detection of open reading frames ("ORFs") and predictions of gene function difficult. Computational methods exist for predicting coding regions in eukaryotic genomes. Gene prediction programs such as GRAIL and GRAIL II, Uberbacher *et al.*, *Proc. Natl. Acad. Sci. USA* 88(24):11261-5 (1991); Xu *et al.*, *Genet. Eng.* 16:241-53 (1994); Uberbacher *et al.*, *Methods Enzymol.* 266:259-81 (1996); GENEFINDER, Solovyev *et al.*, *Nucl. Acids. Res.* 22:5156-63 (1994); Solovyev *et al.*, *Ismb* 5:294-302 (1997); and GEN**[E]**SCAN, Burge *et al.*, *J. Mol. Biol.* 268:78-94 (1997), predict many putative genes without known homology or function. Such programs are known, however, to give high false positive rates. Burset *et al.*, *Genomics* 34:353-367 (1996). Using a consensus obtained by a plurality of such programs is known to increase the reliability of calling exons from genomic sequence. Ansari-Lari *et al.*, Genome Res. 8(1):29-40 (1998).

(2) Herewith, a marked up copy of the paragraph replacing that which begins at line 30 of page 6 and ends at line 9 of page 7:

It is common for microarrays to be derived from cDNA/EST libraries, either from those previously described in the literature, such as those from the I.M.A.G.E. consortium, Lennon *et al.*, "The I.M.A.G.E. Consortium: an Integrated Molecular Analysis of Genomes and Their Expression, *Genomics* 33(1):151-2 (1996), or from the construction of "problem specific" libraries targeted at a particular biological question, R.S. Thomas *et al.*, [*Cancer Res.* (in press)] ***Toxicologist* 54:68-69 (2000)**. Such microarrays by definition can measure expression only of those genes found in EST libraries, and thus have not been useful as probes for genes discovered solely by genomic sequencing.

(3) Herewith, a marked up copy of the paragraph replacing that which begins at line 33 of page 23 and ends at line 8 of page 25:

As discussed below, and further described in detail in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications [filed January 29, 2001 (attorney docket nos. PB 0004 WO 1, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart"; PB 0004 WO 2, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human

brain"; PB 0004 WO 3, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human adult liver"; PB 0004 WO 4, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver"; PB 0004 WO 5, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung"; PB 0004 WO 6, "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow"; PB 0004 WO 7, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human placenta"; PB 0004 WO 8, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in BT 474 cells"; PB 0004 WO 9, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in HBL 100 cells"; PB 0004 WO 10, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in Hela cells")] PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668; PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entirety, we have used the methods and apparatus of the present invention to identify more than 15,000 exons in human genomic sequence whose expression we have confirmed in at least one human tissue or cell type. Fully two-thirds of the exons belong to genes that were not at the time of our discovery represented in existing public expression (EST, cDNA) databases, making the methods and

apparatus of the present invention extremely powerful tools for novel gene discovery.

(4) Herewith, a marked up copy of the paragraph replacing that which begins at line 28 of page 76 and ends at line 3 of page 77:

> One third of the probe sequences (as amplified) produced an exact match (BLAST Expect ("[E]*E*") values less than [1 e^{-100}] **1e-100 (1 x 10^{-100})**) to either an EST (20% of sequences) or a known mRNA (13% of sequences). A further 22% of the probe sequences showed some homology to a known EST or mRNA (BLAST [E]***E*** values from [1 e^{-5}] **1e-5 (1 x 10^{-5})** to [1 e^{-99}] **1e-99 (1 x 10^{-99})**). The remaining 45% of the probe sequences showed no significant sequence homology to any expressed, or potentially expressed, sequences present in public databases.

(5) Herewith a marked-up copy of the paragraph replacing that which begins at line 26 of page 80 and ends at line 10 of page 81:

> FIG. 7A is a matrix presenting the expression of all verified sequences that showed signal intensity greater than 3 in at least one tissue. Each clone is represented by a column in the matrix. Each of the 10 tissues assayed is represented by a separate row in the matrix, and relative expression (expression ratio) of a clone in that tissue is indicated at the respective node

by intensity of green shading, with the intensity legend shown in panel B. The top row of the matrix ("EST Hit") contains "bioinformatic" rather than "physical" expression data – that is, presents the results returned by query of EST, NR and SwissProt databases using the probe sequence. The legend for "bioinformatic expression" (*i.e.*, degree of homology returned) is presented in panel C. Briefly, white is known, black is novel, with gray depicting nonidentical with significant homology (white: [E] ***E*** values < [$1e^{-100}$] **1e-100 (1×10^{-100})**; gray: [E] ***E*** values from [$1e^{-5}$] **1e-5** (1×10^{-5}) to [$1e^{-99}$] **1e-99** (1×10^{-99}); black: [E] ***E*** values > 1e-5 (1×10^{-5})**)**.

(6) Herewith a marked-up copy of the paragraph replacing that at lines 7 - 16 of page 82:

FIG. 8 shows in dashed line the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect [("E")] **("*E*")** value of greater than [$1e^{-30}$] **1e-30** (1×10^{-30}) (designated "unknown") upon query of existing EST, NR and SwissProt databases, and shows [in blue] **solid line** the normalized Cy3 signal intensity for all sequence-verified products with a BLAST Expect value of less than [$1e^{-30}$] **1e-30 (1×10^{-30})** ("known"). Note that biological background noise has an averaged normalized Cy3 signal intensity of 0.2.

(7) Herewith a marked-up copy of the paragraph replacing that which begins at line 18 of page 96 and ends at line 22 of page 97:

> Using this threshold, we identified over 15,000 single exon probes that produce significant signal in one or more of ten tested tissues/cell types. The exact structures of these single exon probes are clearly presented in the SEQUENCE LISTINGs included in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications [filed January 29, 2001 (attorney docket nos. PB 0004 WO 1, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart"; PB 0004 WO 2, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human brain"; PB 0004 WO 3, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human adult liver"; PB 0004 WO 4, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver"; PB 0004 WO 5, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung"; PB 0004 WO 6, "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow"; PB 0004 WO 7, for "Human genome-derived single exon nucleic

acid probes useful for analysis of gene expression in human placenta"; PB 0004 WO 8, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in BT 474 cells"; PB 0004 WO 9, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in HBL 100 cells"; PB 0004 WO 10, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in Hela cells"),] PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668; PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entireties.

(8) Herewith a marked-up copy of the paragraph replacing that which begins at line 12 of page 98 and ends at line 13 of page 99:

The exact structures of these single exon probes are clearly presented in the SEQUENCE LISTINGs included in commonly owned and copending U.S. provisional application nos. 60/207,456, filed May 26, 2000; 60/234,687, filed September 21, 2000; 60/236,359, filed September 27, 2000; in commonly owned and copending U.K. patent application no. 0024263.6, filed October 4, 2000; and in commonly owned and copending PCT applications [filed January 29, 2001 (attorney docket nos. PB 0004 WO 1, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human heart"; PB 0004 WO 2, for "Human genome-

derived single exon nucleic acid probes useful for analysis of gene expression in human brain"; PB 0004 WO 3, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human adult liver"; PB 0004 WO 4, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver"; PB 0004 WO 5, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human lung"; PB 0004 WO 6, "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow"; PB 0004 WO 7, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human placenta"; PB 0004 WO 8, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in BT 474 cells"; PB 0004 WO 9, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in HBL 100 cells"; PB 0004 WO 10, for "Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in Hela cells"),] PCT/US01/00666; PCT/US01/00667; PCT/US01/00664; PCT/US01/00669; PCT/US01/00665; PCT/US01/00668; PCT/US01/00663; PCT/US01/00662; PCT/US01/00661; and PCT/US01/00670, the disclosures of which are incorporated herein by reference in their entireties.